A GENERAL CONVERGENCE ANALYSIS ON INEXACT NEWTON METHOD FOR NONLINEAR INVERSE PROBLEMS

QINIAN JIN

ABSTRACT. We consider the inexact Newton methods

$$x_{n+1}^\delta = x_n^\delta - g_{\alpha_n}\left(F'(x_n^\delta)^* F'(x_n^\delta)\right) F'(x_n^\delta)^* \left(F(x_n^\delta) - y^\delta\right)$$

for solving nonlinear ill-posed inverse problems $F(x) = y$ using the only available noise data y^δ satisfying $\|y^\delta - y\| \le \delta$ with a given small noise level $\delta > 0$. We terminate the iteration by the discrepancy principle

$$\|F(x_{n_\delta}^\delta) - y^\delta\| \le \tau\delta < \|F(x_n^\delta) - y^\delta\|, \qquad 0 \le n < n_\delta$$

with a given number $\tau > 1$. Under certain conditions on $\{\alpha_n\}$ and F, we prove for a large class of spectral filter functions $\{g_\alpha\}$ the convergence of $x_{n_\delta}^\delta$ to a true solution as $\delta \to 0$. Moreover, we derive the order optimal rates of convergence when certain Hölder source conditions hold. Numerical examples are given to test the theoretical results.

1. **Introduction**

In this paper we consider the nonlinear equations

$$(1.1) \qquad F(x) = y,$$

arising from nonlinear inverse problems, where $F : D(F) \subset X \mapsto Y$ is a nonlinear Fréchet differentiable operator between two Hilbert spaces X and Y whose norms and inner products are denoted as $\|\cdot\|$ and $(\cdot, \cdot)$ respectively. We assume that (1.1) has a solution $x^\dagger$ in the domain $D(F)$ of F, i.e. $F(x^\dagger) = y$. We use $F'(x)$ to denote the Fréchet derivative of F at $x \in D(F)$ and $F'(x)^*$ the adjoint of $F'(x)$. A characteristic property of such problems is their ill-posedness in the sense that their solutions do not depend continuously on the data. Since the right hand side y is usually obtained by measurement, the only available data is a noise y^δ satisfying

$$(1.2) \qquad \|y^\delta - y\| \le \delta$$

with a given small noise level $\delta > 0$. Due to the ill-posedness, it is challenging to produce from y^δ a stable approximate solution to $x^\dagger$ and the regularization techniques must be taken into account.

Many regularization methods have been considered for solving (1.1) in the last two decades. Tikhonov regularization is one of the well-known methods that have been studied extensively in the literature. Due to the straightforward implementation, iterative methods are also attractive for solving nonlinear inverse problems. In this paper we will consider a class of inexact Newton methods. To motivate, let x_n^δ be a current iterate. We may approximate $F(x)$ by its linearization around x_n^δ, i.e.

$F(x) \approx F(x_n^\delta) + F'(x_n^\delta)(x - x_n^\delta)$. Thus, instead of (1.1) we have the approximate equation

$$(1.3) \qquad\qquad F'(x_n^\delta)(x - x_n^\delta) = y^\delta - F(x_n^\delta).$$

If $F'(x_n^\delta)$ is invertible, the usual Newton method defines the next iterate by solving (1.3) for x. Computing the exact solution of (1.3) however can be expensive in general even the problem is well-posed. Thus, one might prefer to compute some approximate solution at certain accuracy and use it as the next iterate. This motivates the inexact Newton methods in [2] where for well-posed problems the convergence was carried out when the next computed iterate x_{n+1}^δ satisfies

$$(1.4) \qquad \|F(x_n^\delta) - y^\delta + F'(x_n^\delta)(x_{n+1}^\delta - x_n^\delta)\| \le \mu_n \|F(x_n^\delta) - y^\delta\|$$

at each step with the forcing terms $\mu_n \in (0,1)$ being uniformly bounded below 1. For nonlinear ill-posed inverse problems, $F'(x_n^\delta)$ in general is not invertible and (1.3) usually is ill-posed. Therefore one should use the regularization methods to solve (1.3) approximately. Let $\{g_\alpha\}$ be a family of spectral filter functions. We can apply the linear regularization method defined by $\{g_\alpha\}$ to (1.3) to produce the next iterate. This leads to the following inexact Newton method

$$(1.5) \qquad x_{n+1}^\delta = x_n^\delta - g_{\alpha_n}\left(F'(x_n^\delta)^* F'(x_n^\delta)\right) F'(x_n^\delta)^* \left(F(x_n^\delta) - y^\delta\right),$$

where $x_0^\delta := x_0 \in D(F)$ is an initial guess of $x^\dagger$ and $\{\alpha_n\}$ is a sequence of positive numbers. By taking g_α to be various functions, (1.5) then produces the nonlinear Landweber iteration [5], the Levenberg-Marquardt method [3, 7], the exponential Euler iteration [6], and the first-stage Runge-Kutta type regularization [10].

In this paper we will consider the inexact Newton method (1.5) in a unified way by assuming that $\{\alpha_n\}$ is an a priori given sequence of positive numbers with suitable properties. We will terminate the iteration by the discrepancy principle

$$(1.6) \qquad \|F(x_{n_\delta}^\delta) - y^\delta\| \le \tau\delta < \|F(x_n^\delta) - y^\delta\|, \quad 0 \le n < n_\delta$$

with a given number $\tau > 1$ and consider the approximation property of $x_{n_\delta}^\delta$ to $x^\dagger$ as $\delta \to 0$. For a large class of spectral filter functions $\{g_\alpha\}$ we will establish the convergence of $x_{n_\delta}^\delta$ to $x^\dagger$ as $\delta \to 0$ and derive the order optimal convergence rates for the method defined by (1.5) and (1.6). Our work not only reproduces those known results in [5, 7, 6, 10] but also presents new convergence results and new methods. Furthermore, our convergence analysis provides new insights into the feature of the inexact Newton regularization methods.

In the definition of the inexact Newton method, one may determine the sequence $\{\alpha_n\}$ adaptively during computation. In [3] the Levenberg-Marquardt scheme was considered with $\{\alpha_n\}$ chosen adaptively so that (1.4) holds and the discrepancy principle was used to terminate the iteration. The order optimal convergence rates were derived recently in [4]. The general methods (1.5) with $\{\alpha_n\}$ chosen adaptively to satisfy (1.4) were considered later in [11, 9], but only suboptimal convergence rates were derived in [12] and the convergence analysis is far from complete. The methods of the present paper is essentially different in that the sequence $\{\alpha_n\}$ is given in an a priori way which has the advantage of saving computational work. We hope, however, the work of the present paper can provide better understanding on the methods with $\{\alpha_n\}$ chosen adaptively.

This paper is organized as follows. In Section 2 we first formulate the conditions on $\{\alpha_n\}$, $\{g_\alpha\}$ and F, and state the main results on the convergence and rates

of convergence for the methods defined by (1.5) and (1.6), we then give several examples of iteration methods that fit into the framework (1.5). In Section 3 we prove some crucial inequalities which is frequently used in the convergence analysis. In Section 4 we derive the order optimal convergence rate result when $x_0 - x^\dagger$ satisfies certain source conditions. In Section 5 we show the convergence property without assuming any source conditions on $x_0 - x^\dagger$. Finally in Section 5 we present numerical examples to test the theoretical results.

2. Main results

In order to carry out the convergence analysis on the method defined by (1.5) and (1.6), we need to impose suitable conditions on $\{\alpha_n\}$, $\{g_\alpha\}$ and F. For the sequence $\{\alpha_n\}$ of positive numbers, we set

$$(2.1) \qquad s_{-1} = 0, \qquad s_n := \sum_{j=0}^{n} \frac{1}{\alpha_j}, \qquad n = 0, 1, \cdots.$$

We will assume that there are constants $c_0 > 1$ and $c_1 > 0$ such that

$$(2.2) \qquad \lim_{n\to\infty} s_n = \infty, \quad s_{n+1} \le c_0 s_n \quad \text{and} \quad 0 < \alpha_n \le c_1, \quad n = 0, 1, \cdots.$$

For the spectral filter functions $\{g_\alpha\}$, we will assume the following two conditions, where $\mathbb{C}$ denotes the complex plane.

Assumption 1. *For each $\alpha > 0$, the function*

$$\varphi_\alpha(\lambda) := g_\alpha(\lambda) - \frac{1}{\alpha + \lambda}$$

extends to a complex analytic function defined on a domain $D_\alpha \subset \mathbb{C}$ such that $[0,1] \subset D_\alpha$, and there is a contour $\Gamma_\alpha \subset D_\alpha$ enclosing $[0,1]$ such that

$$(2.3) \qquad |z| \ge \frac{1}{2}\alpha \quad \text{and} \quad \frac{|z| + \lambda}{|z - \lambda|} \le b_0, \qquad \forall z \in \Gamma_\alpha,\, \alpha > 0 \text{ and } \lambda \in [0,1],$$

where b_0 is a constant independent of $\alpha > 0$. Moreover, there is a constant b_1 such that

$$(2.4) \qquad \int_{\Gamma_\alpha} |\varphi_\alpha(z)|\, |dz| \le b_1$$

for all $0 < \alpha \le c_1$.

Assumption 2. *Let $\{\alpha_n\}$ be a sequence of positive numbers, let $\{s_n\}$ be defined by (2.1). There is a constant $b_2 > 0$ such that*

$$(2.5) \qquad 0 \le \lambda^\nu \prod_{k=j}^{n} r_{\alpha_k}(\lambda) \le (s_n - s_{j-1})^{-\nu},$$

$$(2.6) \qquad 0 \le \lambda^\nu g_{\alpha_j}(\lambda) \prod_{k=j+1}^{n} r_{\alpha_k}(\lambda) \le b_2 \frac{1}{\alpha_j}(s_n - s_{j-1})^{-\nu}$$

for $0 \le \nu \le 1$, $0 \le \lambda \le 1$ and $j = 0, 1, \cdots, n$, where $r_\alpha(\lambda) := 1 - \lambda g_\alpha(\lambda)$ is the residual function.

By using the spectral integrals for self-adjoint operators, it follows easily from
(2.3) in Assumption 1 that for any bounded linear operator A with $\|A\| \leq 1$ there
holds

$$(2.7) \qquad \|(zI - A^*A)^{-1}(A^*A)^\nu\| \leq \frac{b_0}{|z|^{1-\nu}}$$

for $z \in \Gamma_\alpha$ and $0 \leq \nu \leq 1$. Moreover, since Assumption 1 implies $\varphi_\alpha(z)$ is analytic
in D_α for each $\alpha > 0$, there holds the Riesz-Dunford formula (see [1])

$$(2.8) \qquad \varphi_\alpha(A^*A) = \frac{1}{2\pi i} \int_{\Gamma_\alpha} \varphi_\alpha(z)(zI - A^*A)^{-1} dz$$

for any linear operator A satisfying $\|A\| \leq 1$.

As a simple consequence of (2.5) in Assumption 2, we have for $0 \leq \nu \leq 1$ and
$\alpha > 0$ that

$$(2.9) \qquad 0 \leq \lambda^\nu(\alpha + \lambda)^{-1} \prod_{k=j+1}^{n} r_{\alpha_k}(\lambda) \leq 2\alpha^{\nu-1}\left(1 + \alpha(s_n - s_j)\right)^{-\nu}$$

for all $0 \leq \lambda \leq 1$ and $j = 0, 1, \cdots, n$, see [8, Lemma 1].

For the nonlinear operator F, we need the following condition which has been
verified in [5] for several nonlinear inverse problems.

Assumption 3. *(a) There exists $K_0 \geq 0$ such that*

$$(2.10) \qquad F'(x) = R(x, \bar{x})F'(\bar{x}) \quad and \quad \|I - R(x, \bar{x})\| \leq K_0 \|x - \bar{x}\|$$

for all $x, \bar{x} \in B_\rho(x^\dagger) \subset D(F)$.
(b) F is properly scaled so that $\|F'(x)\| \leq \min\{1, \sqrt{\alpha_0}\}$ for all $x \in B_\rho(x^\dagger)$.

The condition (a) in Assumption 3 clearly implies that $\|F'(x)\|$ is uniformly
bounded over $B_\rho(x^\dagger)$. Thus, by multiplying (1.1) by a sufficiently small number,
we may assume that F is properly scaled so that condition (b) in Assumption 3 is
satisfied. A direct consequence of Assumption 3 is the inequality

$$\|F(x) - F(x^\dagger) - F'(x^\dagger)(x - x^\dagger)\| \leq \frac{1}{2} K_0 \|x - x^\dagger\| \|F'(x^\dagger)(x - x^\dagger)\|$$

for all $x \in B_\rho(x^\dagger)$, which will be frequently used in the convergence analysis.

Now we are ready to state the first main result concerning the rate of convergence
of $x_{n_\delta}^\delta$ to $x^\dagger$ as $\delta \to 0$ when $e_0 := x_0 - x^\dagger$ satisfies the sourcewise condition

$$(2.11) \qquad x_0 - x^\dagger = (F'(x^\dagger)^* F'(x^\dagger))^\nu \omega$$

for some $0 < \nu \leq 1/2$ and $\omega \in \mathcal{N}(F'(x^\dagger))^\perp \subset X$, where n_δ is the integer determined
by the discrepancy principle (1.6) with $\tau > 1$.

Theorem 2.1. *Let F satisfy Assumptions 3, let $\{g_\alpha\}$ satisfy Assumptions 1 and 2,
and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.2). If $x_0 - x^\dagger$ satisfies
the source condition (2.11) for some $0 < \nu \leq 1/2$ and $\omega \in \mathcal{N}(F'(x^\dagger))^\perp \subset X$ and if
$K_0\|\omega\|$ is suitably small, then*

$$\|x_{n_\delta}^\delta - x^\dagger\| \leq C_\nu \|\omega\|^{1/(1+2\nu)} \delta^{2\nu/(1+2\nu)}$$

*for the integer n_δ determined by the discrepancy principle (1.6) with $\tau > 1$, where
$C_\nu > 0$ is a generic constant independent of δ and $\|\omega\|$.*

Theorem 2.1 shows that the method (1.1) together with the discrepancy principle (1.6) defines an order optimal regularization method for each $0 < \nu \leq 1/2$. This result in particular reproduces the corresponding ones in [5, 7, 6, 10] for various iterative methods even with an improvement by relaxing $\tau > 2$ to $\tau > 1$.

Nevertheless, Theorem 2.1 does not provide the convergence of $x_{n_\delta}^\delta$ to $x^\dagger$ as $\delta \to 0$ if there is no source condition imposed on $x_0 - x^\dagger$. In the next main result we will show the convergence of $x_{n_\delta}^\delta$ to $x^\dagger$ as $\delta \to 0$ if $\{\alpha_n\}$ is a geometric decreasing sequence, i.e.

$$(2.12) \qquad\qquad \alpha_n = \alpha_0 r^n, \qquad n = 0, 1, \cdots$$

for some $\alpha_0 > 0$ and $0 < r < 1$, which is one of the most important cases in applications.

Theorem 2.2. *Let F satisfy Assumptions 3, let $\{g_\alpha\}$ satisfy Assumptions 1 and 2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.12). If $x_0 - x^\dagger \in \mathcal{N}(F'(x^\dagger))^\perp$ and $K_0\|x_0 - x^\dagger\|$ is suitably small, then*

$$\lim_{\delta \to 0} x_{n_\delta}^\delta = x^\dagger$$

for the integer n_δ determined by the discrepancy principle (1.6) with $\tau > 1$.

Theorem 2.2 extends the corresponding result in [7] for the Levenberg-Marquardt method to a general class of methods given by (1.5). The convergence result in Theorem 2.2 still holds if (2.12) is replaced by the condition

$$(2.13) \qquad\qquad d_0 r^n \leq \alpha_n \leq d_1 r^n, \quad n = 0, 1, \cdots$$

for some constants $0 < d_0 \leq d_1 < \infty$ and $0 < r < 1$. In fact, the proof of Theorem 2.2 given in Section 5 requires only $\{\alpha_n\}$ to satisfy (2.2) and (5.1) which can be achieved if $\{\alpha_n\}$ satisfies (2.13). It would be interesting if such a convergence result can be proved for a general sequence $\{\alpha_n\}$ satisfying (2.2) only. This, however, remains open; new technique seems to be explored.

We conclude this section with several examples of the methods (1.5) in which the spectral filter functions $\{g_\alpha\}$ have been shown in [8] to satisfy Assumptions 1 and 2:

(a) We first consider for $\alpha > 0$ the function g_α given by

$$g_\alpha(\lambda) = \frac{(\alpha + \lambda)^N - \alpha^N}{\lambda(\alpha + \lambda)^N}$$

where $N \geq 1$ is a fixed integer. This function arises from the iterated Tikhonov regularization of order N for linear ill-posed problems. The corresponding method (1.5) becomes

$$u_{n,0} = x_n^\delta,$$

$$u_{n,l+1} = u_{n,l} - \left(\alpha_n I + F'(x_n^\delta)^* F'(x_n^\delta)\right)^{-1} F'(x_n^\delta)^* \left(F(x_n^\delta) - y^\delta - F'(x_n^\delta)(x_n^\delta - u_{n,l})\right),$$
$$l = 0, \cdots, N - 1,$$

$$x_{n+1}^\delta = u_{n,N}.$$

When $N = 1$, this is the Levenberg-Marquardt method (see [3, 7]).

(b) We consider the method (1.5) with g_α given by

$$g_\alpha(\lambda) = \frac{1}{\lambda}\left(1 - e^{-\lambda/\alpha}\right)$$

which arises from the asymptotic regularization for linear ill-posed problems. In this method, the iterative sequence $\{x_n^\delta\}$ is equivalently defined as $x_{n+1}^\delta := x(1/\alpha_n)$, where $x(t)$ is the unique solution of the initial value problem

$$\frac{d}{dt}x(t) = F'(x_n^\delta)^* \left(y^\delta - F(x_n^\delta) + F'(x_n^\delta)(x_n^\delta - x(t))\right), \quad t > 0,$$
$$x(0) = x_n^\delta.$$

This is the so called exponential Euler iteration considered in [6].

(c) For $0 < \alpha \le 1$ consider the function g_α given by

$$(2.14) \qquad g_\alpha(\lambda) = \sum_{l=0}^{[1/\alpha]-1} (1 - \lambda)^l = \frac{1 - (1 - \lambda)^{[1/\alpha]}}{\lambda}$$

which arises from the linear Landweber iteration, where $[1/\alpha]$ denotes the largest integer not greater than $1/\alpha$. The method (1.5) then becomes

$$u_{n,0} = x_n^\delta,$$
$$u_{n,l+1} = u_{n,l} - F'(x_n^\delta)^* \left(F(x_n^\delta) - y^\delta - F'(x_n^\delta)(x_n^\delta - u_{n,l})\right), \quad 0 \le l \le [1/\alpha_n] - 1,$$
$$x_{n+1}^\delta = u_{n,[1/\alpha_n]}.$$

When $\alpha_n = 1$ for all n, this method reduces to the Landweber iteration in [5].

(d) For $0 < \alpha \le 1$ consider the function

$$g_\alpha(\lambda) = \sum_{i=1}^{[1/\alpha]} (1 + \lambda)^{-i} = \frac{1 - (1 + \lambda)^{-[1/\alpha]}}{\lambda}$$

arising from the Lardy method for linear inverse problems. Then the method (1.5) becomes

$$u_{n,0} = x_n^\delta,$$
$$u_{n,l+1} = u_{n,l} - \left(I + F'(x_n^\delta)^* F'(x_n^\delta)\right)^{-1} F'(x_n^\delta)^* \left(F(x_n^\delta) - y^\delta - F'(x_n^\delta)(x_n^\delta - u_{n,l})\right),$$
$$l = 0, \cdots, [1/\alpha_n] - 1,$$
$$x_{n+1}^\delta = u_{n,[1/\alpha_n]}.$$

When $\alpha_n = 1$ for all n, this is the so called first-stage Runge-Kutta type regularization considered in [10].

3. Some crucial inequalities

The following consequence of the above assumptions on F and $\{g_\alpha\}$ plays a crucial role in the convergence analysis.

Lemma 3.1. *Let $\{g_\alpha\}$ satisfy Assumptions 1 and 2, let F satisfy Assumption 3, and let $\{\alpha_n\}$ be a sequence of positive numbers. Let $T = F'(x^\dagger)$ and for any $x \in B_\rho(x^\dagger)$ let $T_x = F'(x)$. Let $0 \le a \le 1/2$. Then for $0 \le b \le 1/2 + a$ there holds*

$$(T^*T)^a \prod_{k=j+1}^{n} r_{\alpha_k}(T^*T) \left[g_{\alpha_j}(T^*T)T^* - g_{\alpha_j}(T_x^*T_x)T_x^*\right] = (T^*T)^b S_j$$

for some bounded linear operator $S_j : Y \to X$ satisfying [1]

$$\|S_j\| \lesssim \frac{1}{\alpha_j}(s_n - s_{j-1})^{-1/2-a+b}K_0\|x - x^\dagger\|,$$

where $j = 0, 1, \cdots, n$.

Proof. Let $\eta_\alpha(\lambda) = (\alpha + \lambda)^{-1}$ and $\varphi_\alpha(\lambda) = g_\alpha(\lambda) - (\alpha + \lambda)^{-1}$. We can write

$$(T^*T)^a \prod_{k=j+1}^{n} r_{\alpha_k}(T^*T)\left[g_{\alpha_j}(T^*T)T^* - g_{\alpha_j}(T_x^*T_x)T_x^*\right] = J_1 + J_2 + J_3,$$

where

$$J_1 := (T^*T)^a \prod_{k=j+1}^{n} r_{\alpha_k}(T^*T)g_{\alpha_j}(T^*T)[T^* - T_x^*],$$

$$J_2 := (T^*T)^a \prod_{k=j+1}^{n} r_{\alpha_k}(T^*T)\left[\eta_{\alpha_j}(T^*T) - \eta_{\alpha_j}(T_x^*T_x)\right]T_x^*,$$

$$J_3 := (T^*T)^a \prod_{k=j+1}^{n} r_{\alpha_k}(T^*T)\left[\varphi_{\alpha_j}(T^*T) - \varphi_{\alpha_j}(T_x^*T_x)\right]T_x^*.$$

It suffices to show that for each J_l there holds $J_l = (T^*T)^\nu S_l$ for some bounded linear operator $S_l : Y \to X$ satisfying the desired estimate. We will use the polar decomposition for linear operators which implies that $T^* = (T^*T)^{1/2}U$ for some partial isometry $U : Y \to X$.

By using Assumption 3 we have $T^* - T_x^* = T^*(I - R_x)^*$, where $R_x := R(x, x^\dagger)$. This together with the polar decomposition on T^* gives

(3.1) $$T^* - T_x^* = (T^*T)^{1/2}U(I - R_x)^*.$$

Consequently we can write $J_1 = (T^*T)^b S_1$ with

$$S_1 = (T^*T)^{1/2+a-b}g_{\alpha_j}(T^*T) \prod_{k=j+1}^{n} r_{\alpha_k}(T^*T)U(I - R_x)^*.$$

Since $0 \leq 1/2 + a - b \leq 1$, it follows from Assumption 2 that

$$\|S_1\| \leq \sup_{0 \leq \lambda \leq 1} \left(\lambda^{1/2+a-b}g_{\alpha_j}(\lambda) \prod_{k=j+1}^{n} r_{\alpha_k}(\lambda)\right) \|I - R_x\|$$

$$\lesssim \frac{1}{\alpha_j}(s_n - s_{j-1})^{-1/2-a+b}K_0\|x - x^\dagger\|.$$

This shows the desired conclusion on J_1.

Next we consider J_2. Note that

$$\eta_{\alpha_j}(T^*T) - \eta_{\alpha_j}(T_x^*T_x) = (\alpha_j I + T^*T)^{-1}T^*(T_x - T)(\alpha_j I + T_x^*T_x)^{-1}$$
$$+ (\alpha_j I + T^*T)^{-1}(T_x^* - T^*)T_x(\alpha_j I + T_x^*T_x)^{-1}.$$

---

[1] Throughout this paper we will always use C to denote a generic constant independent of δ and n. We will also use the convention $\Phi \lesssim \Psi$ to mean that $\Phi \leq C\Psi$ for some generic constant C.

Plugging this formula into the expression of J_2, and using the polar decomposition on T^* and the identity (3.1), we have $J_2 = (T^*T)^b S_2$, where

$$S_2 = \prod_{k=j+1}^{n} r_{\alpha_k}(T^*T)(\alpha_j I + T^*T)^{-1}(T^*T)^{1/2+a-b}U(T_x - T)(\alpha_j I + T_x^* T_x)^{-1}T_x^*$$

$$+ \prod_{k=j+1}^{n} r_{\alpha_k}(T^*T)(\alpha_j I + T^*T)^{-1}(T^*T)^{1/2+a-b}U(R_x - I)^*T_x T_x^*(\alpha_j I + T_x T_x^*)^{-1}.$$

With the help of Assumption 3 we have

$$\|(T_x - T)(\alpha_j I + T_x^* T_x)^{-1}T_x^*\| \le K_0 \|x - x^\dagger\|.$$

Therefore, it follows from (2.9) that

$$\|S_2\| \le \sup_{0 \le \lambda \le 1} \left(\lambda^{1/2+a-b}(\alpha_j + \lambda)^{-1} \prod_{k=j+1}^{n} r_{\alpha_k}(\lambda) \right) \|(T_x - T)(\alpha_j I + T_x^* T_x)^{-1}T_x^*\|$$

$$+ \sup_{0 \le \lambda \le 1} \left(\lambda^{1/2+a-b}(\alpha_j + \lambda)^{-1} \prod_{k=j+1}^{n} r_{\alpha_k}(\lambda) \right) \|(R_x - I)^*T_x T_x^*(\alpha_j I + T_x^* T_x)^{-1}\|$$

$$\lesssim \alpha_j^{a-b-1/2}\left(1 + \alpha_j(s_n - s_j)\right)^{-1/2-a+b} K_0 \|x - x^\dagger\|$$

$$= \frac{1}{\alpha_j}(s_n - s_{j-1})^{-1/2-a+b}K_0\|x - x^\dagger\|.$$

It remains to consider J_3. Since Assumption 1 implies that $\varphi_{\alpha_j}(z)$ is analytic in D_{α_j}, we have from the Riesz-Dunford formula (2.8) that

$$(3.2) \qquad\qquad J_3 = \frac{1}{2\pi i}\int_{\Gamma_{\alpha_j}} \varphi_{\alpha_j}(z)L_j(z)dz,$$

where

$$L_j(z) := (T^*T)^a \prod_{k=j+1}^{n} r_{\alpha_k}(T^*T)\left[(zI - T^*T)^{-1} - (zI - T_x^* T_x)^{-1}\right]T_x^*.$$

Using the decomposition

$$(zI - T^*T)^{-1} - (zI - T_x^* T_x)^{-1} = (zI - T^*T)^{-1}T^*(T - T_x)(zI - T_x^* T_x)^{-1}$$

$$+ (zI - T^*T)^{-1}(T^* - T_x^*)T_x(zI - T_x^* T_x)^{-1}$$

together with the polar decomposition on T^* and the identity (3.1), we obtain $L_j(z) = (T^*T)^b \tilde{L}_j(z)$, where

$$\tilde{L}_j(z) = \prod_{k=j+1}^{n} r_{\alpha_k}(T^*T)(zI - T^*T)^{-1}(T^*T)^{1/2+a-b}(T - T_x)(zI - T_x^* T_x)^{-1}T_x^*$$

$$+ \prod_{k=j+1}^{n} r_{\alpha_k}(T^*T)(zI - T^*T)^{-1}(T^*T)^{1/2+a-b}U(I - R_x)^*T_x T_x^*(zI - T_x T_x^*)^{-1}.$$

Combining with (3.2) gives $J_3 = (T^*T)^b S_3$, where

$$S_3 = \frac{1}{2\pi i}\int_{\Gamma_{\alpha_j}} \varphi_{\alpha_j}(z)\tilde{L}_j(z)dz.$$

We need to estimate $\|S_3\|$. We first estimate $\tilde{L}_j(z)$ for $z \in \Gamma_{\alpha_j}$. With the help of Assumption 3 and (2.7), we have

$$\|(T - T_x)(zI - T_x^* T_x)^{-1} T_x^*\| \lesssim K_0 \|x - x^\dagger\|.$$

Since $|z| \geq \alpha_j/2$ and $|z - \lambda|^{-1} \leq b_0(|z| + \lambda)^{-1}$ for $z \in \Gamma_{\alpha_j}$, we have from (2.9) that

$$
\begin{aligned}
\|\tilde{L}_j(z)\| &\lesssim \sup_{0 \leq \lambda \leq 1} \left(\lambda^{1/2 + a - b} |z - \lambda|^{-1} \prod_{k=j+1}^n r_{\alpha_k}(\lambda) \right) K_0 \|x - x^\dagger\| \\
&\lesssim \sup_{0 \leq \lambda \leq 1} \left(\lambda^{1/2 + a - b} (|z| + \lambda)^{-1} \prod_{k=j+1}^n r_{\alpha_k}(\lambda) \right) K_0 \|x - x^\dagger\| \\
&\lesssim |z|^{a - b - 1/2} \left(1 + (s_n - s_j)|z| \right)^{-1/2 - a + b} K_0 \|x - x^\dagger\| \\
&\lesssim \alpha_j^{a - b - 1/2} \left(1 + (s_n - s_j)\alpha_j \right)^{-1/2 - a + b} K_0 \|x - x^\dagger\| \\
&= \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1/2 - a + b} K_0 \|x - x^\dagger\|.
\end{aligned}
$$

Therefore, it follows from Assumption 1 that

$$
\begin{aligned}
\|S_3\| &\lesssim \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1/2 - a + b} K_0 \|x - x^\dagger\| \int_{\Gamma_{\alpha_j}} |\varphi_{\alpha_j}(z)| |dz| \\
&\lesssim \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1/2 - a + b} K_0 \|x - x^\dagger\|.
\end{aligned}
$$

The proof is therefore complete. $\qquad\square$

In the proof of Theorem 2.2 we will also need the following inequality which can be obtained by essentially the same argument in the proof of Lemma 3.1.

Lemma 3.2. *Let $\{g_\alpha\}$ satisfy Assumptions 1 and 2, let F satisfy Assumption 3, and let $\{\alpha_n\}$ be a sequence of positive numbers. Let $T = F'(x^\dagger)$ and for any $x \in B_\rho(x^\dagger)$ let $T_x = F'(x)$. Then for $0 \leq \mu \leq 1/2$ there holds*

$$
\left\| (T^* T)^\mu \prod_{k=j+1}^n r_{\alpha_k}(T^* T) \left[g_{\alpha_j}(T_x^* T_x) T_x^* - g_{\alpha_j}(T_{\bar{x}}^* T_{\bar{x}}) T_{\bar{x}}^* \right] \right\|
$$

$$
\lesssim \frac{1}{\alpha_j} (s_n - s_{j-1})^{-\mu - 1/2} \left(1 + K_0 \|x - x^\dagger\| \right) K_0 \|x - \bar{x}\|
$$

for all $x, \bar{x} \in B_\rho(x^\dagger)$, where $j = 0, 1, \cdots, n$.

4. Rates of convergence: proof of Theorem 2.1

We begin with the following lemma which follows from [4, Lemma 4.3] and its proof; a simplified argument can be found in [8].

Lemma 4.1. *Let $\{\alpha_n\}$ be a sequence of positive numbers satisfying $\alpha_n \leq c_1$, and let s_n be defined by (2.1). Let $p \geq 0$ and $q \geq 0$ be two numbers. Then we have*

$$
\sum_{j=0}^n \frac{1}{\alpha_j} (s_n - s_{j-1})^{-p} s_j^{-q} \leq C_0 s_n^{1 - p - q} \begin{cases} 1, & \max\{p, q\} < 1, \\ \log(1 + s_n), & \max\{p, q\} = 1, \\ s_n^{\max\{p,q\} - 1}, & \max\{p, q\} > 1, \end{cases}
$$

where C_0 is a constant depending only on c_1, p and q.

In order to derive the necessary estimates on $x_n^\delta - x^\dagger$, we need some useful identities. For simplicity of presentation, in this section we set

$$e_n^\delta := x_n^\delta - x^\dagger, \quad T := F'(x^\dagger) \quad \text{and} \quad T_n := F'(x_n^\delta).$$

It follows from (1.5) that

$$e_{n+1}^\delta = e_n^\delta - g_{\alpha_n}(T_n^* T_n) T_n^* (F(x_n^\delta) - y^\delta).$$

Let

$$u_n := F(x_n^\delta) - y - T(x_n^\delta - x^\dagger).$$

Then we can write

$$
\begin{aligned}
e_{n+1}^\delta = {}& e_n^\delta - g_{\alpha_n}(T^*T)T^*(F(x_n^\delta) - y^\delta) \\
& - \left[g_{\alpha_n}(T_n^* T_n)T_n^* - g_{\alpha_n}(T^*T)T^* \right](F(x_n^\delta) - y^\delta) \\
= {}& r_{\alpha_n}(T^*T)e_n^\delta - g_{\alpha_n}(T^*T)T^*(y - y^\delta + u_n) \\
& - \left[g_{\alpha_n}(T_n^* T_n)T_n^* - g_{\alpha_n}(T^*T)T^* \right](F(x_n^\delta) - y^\delta).
\end{aligned}
$$
(4.1)

By telescoping (4.1) we can obtain

$$
\begin{aligned}
e_{n+1}^\delta = {}& \prod_{j=0}^n r_{\alpha_j}(T^*T)e_0 - \sum_{j=0}^n \prod_{k=j+1}^n r_{\alpha_k}(T^*T)g_{\alpha_j}(T^*T)T^*(y - y^\delta + u_j) \\
& - \sum_{j=0}^n \prod_{k=j+1}^n r_{\alpha_k}(T^*T)\left[g_{\alpha_j}(T_j^* T_j)T_j^* - g_{\alpha_j}(T^*T)T^* \right](F(x_j^\delta) - y^\delta).
\end{aligned}
$$
(4.2)

By multiplying (4.2) by $T := F'(x^\dagger)$ and noting that

$$
(4.3) \qquad I - \sum_{j=0}^n \prod_{k=j+1}^n r_{\alpha_k}(TT^*)g_{\alpha_j}(TT^*)TT^* = \prod_{j=0}^n r_{\alpha_j}(TT^*),
$$

we can obtain

$$
\begin{aligned}
& Te_{n+1}^\delta - y^\delta + y \\
& = T\prod_{j=0}^n r_{\alpha_j}(T^*T)e_0 + \prod_{j=0}^n r_{\alpha_j}(TT^*)(y - y^\delta) - \sum_{j=0}^n \prod_{k=j+1}^n r_{\alpha_k}(TT^*)g_{\alpha_j}(TT^*)TT^* u_j \\
& \quad - \sum_{j=0}^n T\prod_{k=j+1}^n r_{\alpha_k}(T^*T)\left[g_{\alpha_j}(T_j^* T_j)T_j^* - g_{\alpha_j}(T^*T)T^* \right](F(x_j^\delta) - y^\delta).
\end{aligned}
$$
(4.4)

Based on (4.2) and (4.4) we will prove Theorem 2.1 concerning the order optimal convergence rate of $x_{n_\delta}^\delta$ to $x^\dagger$ when $e_0 := x_0 - x^\dagger$ satisfies the source condition (2.11) for some $0 < \nu \le 1/2$ and $\omega \in \mathcal{N}(F'(x^\dagger))^\perp \subset X$. We will first derive the crucial estimates on $\|e_n^\delta\|$ and $\|Te_n^\delta\|$. To this end, we introduce the integer $\tilde{n}_\delta$ satisfying

$$
(4.5) \qquad s_{\tilde{n}_\delta}^{-\nu-1/2} \le \frac{(\tau-1)\delta}{2c_0\|\omega\|} < s_n^{-\nu-1/2}, \qquad 0 \le n < \tilde{n}_\delta,
$$

where $c_0 > 1$ is the constant appearing in (2.2). Such $\tilde{n}_\delta$ is well-defined since $s_n \to \infty$ as $n \to \infty$.

Proposition 4.1. *Let F satisfy Assumptions 3, let $\{g_\alpha\}$ satisfy Assumptions 1 and 2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.2). If $x_0 - x^\dagger$ satisfies (2.11) for some $0 < \nu \le 1/2$ and $\omega \in \mathcal{N}(F'(x^\dagger))^\perp \subset X$ and if $K_0\|\omega\|$ is suitably small, then there exists a generic constant $C_* > 0$ such that*

$$(4.6) \qquad \|e_n^\delta\| \le C_* s_n^{-\nu}\|\omega\| \qquad and \qquad \|Te_n^\delta\| \le C_* s_n^{-\nu-1/2}\|\omega\|$$

and

$$(4.7) \qquad \|Te_n^\delta - y^\delta + y\| \le (c_0 + C_* K_0\|\omega\|)s_n^{-\nu-1/2}\|\omega\| + \delta$$

for all $0 \le n \le \tilde{n}_\delta$.

Proof. We will show (4.6) by induction. By using (2.11) and $\|T\| \le \sqrt{\alpha_0}$ it is easy to see that (4.6) for $n = 0$ holds if $C_* \ge 1$. Next we assume that (4.6) holds for all $0 \le n \le l$ for some $l < \tilde{n}_\delta$ and show (4.6) holds for $n = l + 1$.

With the help of (2.11) we can derive from (4.2) that

$$\|e_{l+1}^\delta\| \le \left\|\prod_{j=0}^l r_{\alpha_j}(T^*T)(T^*T)^\nu\omega\right\| + \left\|\sum_{j=0}^l \prod_{k=j+1}^l r_{\alpha_k}(T^*T)g_{\alpha_j}(T^*T)T^*(y - y^\delta + u_j)\right\|$$

$$+ \left\|\sum_{j=0}^l \prod_{k=j+1}^l r_{\alpha_k}(T^*T)\left[g_{\alpha_j}(T_j^*T_j)T_j^* - g_{\alpha_j}(T^*T)T^*\right](F(x_j^\delta) - y^\delta)\right\|.$$

Thus we may use Assumption 2 and Lemma 3.1 with $a = b = 0$ to conclude

$$\|e_{l+1}^\delta\| \le s_l^{-\nu}\|\omega\| + b_2 \sum_{j=0}^l \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1/2}(\delta + \|u_j\|)$$

$$(4.8) \qquad + C\sum_{j=0}^l \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1/2}K_0\|e_j\|\|F(x_j^\delta) - y^\delta\|.$$

Moreover, by using (2.11), Assumption 2 and Lemma 3.1 with $a = 1/2$ and $b = 0$, we have from (4.4) that

$$\|Te_{l+1}^\delta - y^\delta + y\| \le s_l^{-\nu-1/2}\|\omega\| + \delta + b_2\sum_{j=0}^l \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1}\|u_j\|$$

$$(4.9) \qquad + C\sum_{j=0}^l \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1}K_0\|e_j\|\|F(x_j^\delta) - y^\delta\|.$$

With the help of Assumption 3 and the induction hypotheses, it follows for all $0 \le j \le l$ that

$$(4.10) \qquad \|u_j\| \le K_0\|e_j^\delta\|\|Te_j^\delta\| \lesssim K_0\|\omega\|^2 s_j^{-2\nu-1/2}.$$

By using the fact

$$(4.11) \qquad \delta \le \frac{2c_0}{\tau - 1}\|\omega\|s_j^{-\nu-1/2}, \qquad 0 \le j \le l$$

and the induction hypotheses we have

$$(4.12) \qquad \|F(x_j^\delta) - y^\delta\| \le \delta + \|Te_j^\delta\| + \|u_j\| \lesssim \|\omega\|s_j^{-\nu-1/2}.$$

In view of the estimates (4.10), (4.12), the induction hypothesis on $\|e_j\|$ and the inequality

$$(4.13) \qquad \sum_{j=0}^{l} \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} \le c_2 s_l^{1/2}$$

for some generic constant c_2, which follows from Lemma 4.1, we have from (4.8) and (4.9) that

$$\|e_{l+1}^{\delta}\| \le \|\omega\| s_l^{-\nu} + c_2 s_l^{1/2} \delta + CK_0 \|\omega\|^2 \sum_{j=0}^{l} \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} s_j^{-2\nu-1/2}$$

and

$$\|Te_{l+1}^{\delta} - y^{\delta} + y\| \le \|\omega\| s_l^{-\nu-1/2} + \delta + CK_0 \|\omega\|^2 \sum_{j=0}^{l} \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} s_j^{-2\nu-1/2}.$$

With the help of Lemma 4.1, $\nu > 0$, (4.11) and (2.2) we have

$$\|e_{l+1}^{\delta}\| \le \left(1 + \frac{2}{\tau-1} c_0 c_2 + CK_0 \|\omega\|\right) \|\omega\| s_l^{-\nu}$$

and

$$\|Te_{l+1}^{\delta} - y^{\delta} + y\| \le \delta + (1 + CK_0\|\omega\|) \|\omega\| s_l^{-\nu-1/2}$$

$$(4.14) \qquad\qquad\qquad \le \delta + (c_0 + CK_0\|\omega\|) \|\omega\| s_{l+1}^{-\nu-1/2}.$$

Consequently $\|e_{l+1}^{\delta}\| \le C_* \|\omega\| s_{l+1}^{-\nu}$ if $C_* \ge 2 + \frac{2}{\tau-1} c_0 c_2$ and $K_0 \|\omega\|$ is suitably small. Moreover, from (4.14), (4.11) and (2.2) we also have

$$\|Te_{l+1}^{\delta}\| \le 2\delta + (c_0 + CK_0\|\omega\|) \|\omega\| s_{l+1}^{-\nu-1/2}$$

$$\le \left(\frac{4c_0^2}{\tau-1} + c_0 + CK_0\|\omega\|\right) \|\omega\| s_{l+1}^{-\nu-1/2}$$

$$\le C_* \|\omega\| s_{l+1}^{-\nu-1/2}$$

if $C_* \ge 2c_0 + \frac{4c_0^2}{\tau-1}$ and $K_0\|\omega\|$ is suitably small. We therefore complete the proof of (4.6). In the meanwhile, (4.14) gives the proof of (4.7). $\qquad\square$

From Proposition 4.1 it follows that $x_n \in B_\rho(x^\dagger)$ for $0 \le n \le \tilde{n}_\delta$ if $\|\omega\|$ is sufficiently small. Furthermore, from (4.10) and (4.12) we have

$$(4.15) \qquad\qquad \|F(x_n^{\delta}) - y - Te_n^{\delta}\| \lesssim K_0 \|\omega\|^2 s_n^{-2\nu-1/2}$$

and

$$(4.16) \qquad\qquad \|F(x_n^{\delta}) - y^{\delta}\| \lesssim \|\omega\| s_n^{-\nu-1/2}$$

for $0 \le n \le \tilde{n}_\delta$.

In the following we will show that $n_\delta \le \tilde{n}_\delta$ for the integer n_δ defined by (1.6) with $\tau > 1$. Consequently, the method given by (1.5) and (1.6) is well-defined.

Lemma 4.2. *Let all the conditions in Proposition 4.1 hold. Let $\tau > 1$ be a given number. If $x_0 - x^\dagger$ satisfies (2.11) for some $0 < \nu \le 1/2$ and $\omega \in \mathcal{N}(F'(x^\dagger))^\perp \subset X$ and if $K_0\|\omega\|$ is suitably small, then the discrepancy principle (1.6) defines a finite integer n_δ satisfying $n_\delta \le \tilde{n}_\delta$.*

Proof. From Proposition 4.1, (4.15) and $\nu > 0$ it follows for $0 \leq n \leq \tilde{n}_\delta$ that

$$\|F(x_n^\delta) - y^\delta\| \leq \|F(x_n^\delta) - y - Te_n^\delta\| + \|Te_n^\delta - y^\delta + y\|$$
$$\leq CK_0\|\omega\|^2 s_n^{-2\nu-1/2} + (c_0 + CK_0\|\omega\|) s_n^{-\nu-1/2}\|\omega\| + \delta$$
$$\leq (c_0 + CK_0\|\omega\|) s_n^{-\nu-1/2}\|\omega\| + \delta.$$

By setting $n = \tilde{n}_\delta$ in the above inequality and using the definition of $\tilde{n}_\delta$ we obtain

$$\|F(x_{\tilde{n}_\delta}^\delta) - y^\delta\| \leq \left(1 + \frac{\tau - 1}{2} + CK_0\|\omega\|\right)\delta \leq \tau\delta$$

if $K_0\|\omega\|$ is suitably small. According to the definition of n_δ we have $n_\delta \leq \tilde{n}_\delta$. $\qquad\square$

4.1. **Completion of the proof of Theorem 2.1.** From (4.2), the source condition (2.11), the polar decomposition on T^*, and Lemma 3.1 with $a = 0$ and $b = \nu$ it follows that

$$(4.17) \qquad\qquad e_{n+1}^\delta = (T^*T)^\nu w_n,$$

where

$$w_n := \prod_{j=0}^n r_{\alpha_j}(T^*T)\omega - \sum_{j=0}^n S_j(F(x_j^\delta) - y^\delta)$$
$$- \sum_{j=0}^n \prod_{k=j+1}^n r_{\alpha_k}(T^*T)g_{\alpha_j}(T^*T)(T^*T)^{1/2-\nu}U(y - y^\delta + u_j).$$

With the help of Assumption 1 and Lemma 3.1 we have

$$\|w_n\| \lesssim \|\omega\| + \sum_{j=0}^n \frac{1}{\alpha_j}(s_n - s_{j-1})^{-1/2+\nu}K_0\|e_j\|\|F(x_j^\delta) - y^\delta\|$$
$$+ \sum_{j=0}^n \frac{1}{\alpha_j}(s_n - s_{j-1})^{-1/2+\nu}(\delta + \|u_j\|).$$

In view of (4.15), (4.16), Proposition 4.1, Lemma 4.1 and (4.5) it yields for $0 \leq n < \tilde{n}_\delta$ that

$$\|w_n\| \lesssim \|\omega\| + \delta\sum_{j=0}^n \frac{1}{\alpha_j}(s_n - s_{j-1})^{-1/2+\nu}$$
$$+ K_0\|\omega\|^2 \sum_{j=0}^n \frac{1}{\alpha_j}(s_n - s_{j-1})^{-1/2+\nu}s_j^{-2\nu-1/2}$$
$$\lesssim \|\omega\| + s_n^{1/2+\nu}\delta \lesssim \|\omega\|.$$

Since Lemma 4.2 implies that $n_\delta \leq \tilde{n}_\delta$, we have $\|w_{n_\delta-1}\| \lesssim \|\omega\|$. On the other hand, it follows from (4.17), Assumption 3 and the definition of n_δ that

$$\|T(T^*T)^\nu w_{n_\delta-1}\| = \|Te_{n_\delta}\| \lesssim \|F(x_{n_\delta}^\delta) - y\| \lesssim \delta.$$

Therefore, by using (4.17) and the above two estimates, we have from the interpolation inequality that

$$\|e_{n_\delta}^\delta\| \leq \|w_{n_\delta-1}\|^{1/(1+2\nu)}\|T(T^*T)^\nu w_{n_\delta-1}\|^{2\nu/(1+2\nu)}$$
$$\leq C_\nu\|\omega\|^{1/(1+2\nu)}\delta^{2\nu/(1+2\nu)}.$$

This gives the desired estimate.

5. Convergence: proof of Theorem 2.2

In this section we will show Theorem 2.2 concerning the convergence of $x_{n_\delta}^\delta$ to $x^\dagger$ as $\delta \to 0$ without assuming any source conditions on $e_0 := x_0 - x^\dagger$. The sequence $\{\alpha_n\}$ is now given by (2.12). It is easy to see that $1/\alpha_n \le s_n \le 1/((1-r)\alpha_n)$ and

$$(5.1) \qquad \sum_{j=0}^{n} \frac{1}{\alpha_j}(s_n - s_{j-1})^{-1}s_j^{-\mu} \le C_1 s_n^{-\mu}$$

for $0 \le \mu < 1$, where C_1 is a constant depending only on r and μ. We remark that (5.1) may not be true for a general sequence $\{\alpha_n\}$ satisfying (2.2).

We first show that the method given by (1.5) and (1.6) is well-defined. To this end, we introduce the integer $\hat{n}_\delta$ satisfying

$$(5.2) \qquad s_{\hat{n}_\delta}^{-1/2} \le \frac{(\tau-1)\delta}{2c_0\|e_0\|} < s_n^{-1/2}, \qquad 0 \le n < \hat{n}_\delta.$$

Since $s_n \to \infty$ as $n \to \infty$, such $\hat{n}_\delta$ is well-defined.

Lemma 5.1. *Let F satisfy Assumptions 3, let $\{g_\alpha\}$ satisfy Assumptions 1 and 2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.12). Let $\tau > 1$ be a given number. If $K_0\|e_0\|$ is suitably small, then there is a constant C_* such that*

$$(5.3) \qquad \|e_n^\delta\| \le C_*\|e_0\| \quad and \quad \|Te_n^\delta\| \le C_*\|e_0\|s_n^{-1/2}$$

for $0 \le n \le \hat{n}_\delta$, and the discrepancy principle (1.6) defines a finite integer n_δ satisfying $n_\delta \le \hat{n}_\delta$.

Proof. We prove (5.3) by induction. By using $\|T\| \le \sqrt{\alpha_0}$, it is easy to see that (5.9) is true for $n = 0$ if $C_* \ge 1$. Next we assume that (5.9) holds for all $0 \le n \le l$ for some $l < \hat{n}_\delta$, and show that it is also true for $n = l+1$. By a similar argument in the proof of Proposition 4.1 we obtain

$$(5.4) \qquad \|e_{l+1}^\delta\| \le \|e_0\| + c_2 s_l^{1/2}\delta + CK_0\|e_0\|^2 \sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1/2}s_j^{-1/2}$$

and

$$(5.5) \qquad \|Te_{l+1}^\delta - y^\delta + y\| \le \|e_0\|s_l^{-1/2} + \delta + CK_0\|e_0\|^2 \sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1}s_j^{-1/2}.$$

By using (5.2) and Lemma 4.1 we obtain from (5.4) that

$$\|e_{l+1}^\delta\| \le \left(1 + \frac{2}{\tau-1}c_0 c_2 + CK_0\|e_0\|\right)\|e_0\| \le C_*\|e_0\|$$

if $C_* \ge 2 + \frac{2}{\tau-2}c_0 c_2$ and $K_0\|e_0\|$ is suitably small. On the other hand, by using (5.1) with $\mu = 1/2$ and (2.2) we obtain from (5.5) that

$$\|Te_{l+1}^\delta - y^\delta + y\| \le \delta + (1 + CK_0\|e_0\|)\|e_0\|s_l^{-1/2}$$

$$(5.6) \qquad\qquad\qquad \le \delta + (c_0 + CK_0\|e_0\|)\|e_0\|s_{l+1}^{-1/2}.$$

Consequently, we have from (5.2) that

$$\|Te_{l+1}^\delta\| \le \left(\frac{4c_0^2}{\tau-1} + c_0 + CK_0\|e_0\|\right)\|e_0\|s_{l+1}^{-1/2} \le C_*\|e_0\|s_{l+1}^{-1/2}$$

if $C_* \geq 2c_0 + \frac{4c_0^2}{\tau-1}$ and $K_0\|e_0\|$ is suitably small. We thus complete the proof of (5.3).

Note that the above argument in fact shows also that

$$\|Te_n^\delta - y^\delta + y\| \leq \delta + (c_0 + CK_0\|e_0\|)\|e_0\|s_n^{-1/2}, \qquad 0 \leq n \leq \hat{n}_\delta.$$

Thus, by the similar argument in the proof of Lemma 4.2 we can derive $\|F(x_{\hat{n}_\delta}^\delta) - y^\delta\| \leq \tau\delta$ if $K_0\|e_0\|$ is suitably small. According to the definition of n_δ we obtain $n_\delta \leq \hat{n}_\delta$. $\qquad\square$

In the remaining part of this section we will show $x_{n_\delta}^\delta \to x^\dagger$ as $\delta \to 0$. We will achieve this by first considering the noise free iterative sequence $\{x_n\}$ defined by (1.5) with y^δ replaced by y, i.e.

$$(5.7) \qquad x_{n+1} = x_n - g_{\alpha_n}\left(F'(x_n)^* F'(x_n)\right) F'(x_n)^*(F(x_n) - y)$$

and showing that $x_n \to x^\dagger$ as $n \to \infty$. We then derive the stability estimate on $\|x_n^\delta - x_n\|$ for $0 \leq n \leq n_\delta$ together with other related estimates. With the help of the definition of n_δ, we will be able to show the convergence of $x_{n_\delta}^\delta$ to $x^\dagger$ as $\delta \to 0$.

5.1. Convergence of the noise free iteration.
In this subsection we will show the convergence of x_n to $x^\dagger$ as $n \to \infty$. We first show that if $x_0 - x^\dagger \in \mathcal{R}(F'(x^\dagger)^*)$ then $x_n \to x^\dagger$ as $n \to \infty$. We then perturb the initial guess x_0 to be $\hat{x}_0$ such that $\hat{x}_0 - x^\dagger \in \mathcal{R}(F'(x^\dagger)^*)$ and define $\{\hat{x}_n\}$ by

$$(5.8) \qquad \hat{x}_{n+1} = \hat{x}_n - g_{\alpha_n}\left(F'(\hat{x}_n)^* F'(\hat{x}_n)\right) F'(\hat{x}_n)^*(F(\hat{x}_n) - y).$$

Since $x_0 - x^\dagger \in \mathcal{N}(F'(x^\dagger))^\perp = \overline{\mathcal{R}(F'(x^\dagger)^*)}$, such $\hat{x}_0$ can be chosen as close to x_0 as we want. We then show that $\{x_n\}$ is stable relative to the change of x_0. This allows us to derive the convergence of $\{x_n\}$.

We start with several lemmas. We first show that x_n is well-defined for all n and satisfies certain estimates.

Lemma 5.2. *Let F satisfy Assumptions 3, let $\{g_\alpha\}$ satisfy Assumptions 1 and 2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.12). If $K_0\|e_0\|$ is suitably small, then*

$$(5.9) \qquad \|e_n\| \leq 2\|e_0\| \quad and \quad \|Te_n\| \leq 2c_0\|e_0\|s_n^{-1/2}$$

for $n = 0, 1, \cdots$, where $e_n := x_n - x^\dagger$.

Proof. This result can be obtained by the same argument in the proof of Lemma 5.1. $\qquad\square$

Lemma 5.3. *Let F satisfy Assumptions 3, let $\{g_\alpha\}$ satisfy Assumptions 1 and 2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.12). If $e_0 = (T^*T)^{1/4}\omega$ for some $\omega \in \mathcal{N}(T)^\perp \subset X$ and $K_0\|e_0\|$ is suitably small, then*

$$(5.10) \qquad \|e_n\| \leq 2c_0\|\omega\|s_n^{-1/4} \quad and \quad \|Te_n\| \leq 2c_0\|\omega\|s_n^{-3/4}$$

for $n = 0, 1, \cdots$.

Proof. We prove (5.10) by induction. By using $\|T\| \leq \sqrt{\alpha_0}$ and $e_0 = (T^*T)^{1/4}\omega$, it is easy to see that (5.10) is true for $n = 0$. Next we assume that (5.10) holds for

all $0 \le n \le l$, and show that it also holds for $n = l + 1$. By a similar argument in the proof of Proposition 4.1 we obtain

$$\|e_{l+1}\| \le s_l^{-1/4}\|\omega\| + b_2 \sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1/2}\|F(x_j) - y - Te_j\|$$

(5.11)
$$+ C\sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1/2}K_0\|e_j\|\|F(x_j) - y\|$$

and

$$\|Te_{l+1}\| \le s_l^{-3/4}\|\omega\| + b_2 \sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1}\|F(x_j) - y - Te_j\|$$

(5.12)
$$+ C\sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1}K_0\|e_j\|\|F(x_j) - y\|.$$

With the help of Assumption 3, Lemma 5.2 and the induction hypotheses, we have for $0 \le j \le l$ that

$$\|F(x_j) - y - Te_j\| \le K_0\|e_j\|\|Te_j\| \le CK_0\|e_0\|\|\omega\|s_j^{-3/4},$$
$$\|F(x_j) - y\| \le \|Te_j\| + \|F(x_j) - y - Te_j\| \lesssim \|\omega\|s_j^{-3/4}.$$

Therefore, by using Lemma 4.1, we obtain from (5.11) that

$$\|e_{l+1}\| \le s_l^{-1/4}\|\omega\| + CK_0\|e_0\|\|\omega\| \sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1/2}s_j^{-3/4}$$
$$\le (1 + CK_0\|e_0\|)\|\omega\|s_l^{-1/4},$$

while by using (5.1) with $\mu = 3/4$ we obtain

$$\|Te_{l+1}\| \le s_l^{-3/4}\|\omega\| + CK_0\|e_0\|\|\omega\| \sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1}s_j^{-3/4}$$
$$\le (1 + CK_0\|e_0\|)\|\omega\|s_l^{-3/4}.$$

Thus, by using $s_{l+1} \le c_0 s_l$, we obtain for suitably small $K_0\|e_0\|$ that $\|e_{l+1}\| \le 2c_0\|\omega\|s_{l+1}^{-1/4}$ and $\|Te_{l+1}\| \le 2c_0\|\omega\|s_{l+1}^{-3/4}$. The proof is therefore complete. $\square$

We remark that the crucial point in Lemma 5.1 is that it requires only the smallness of $K_0\|e_0\|$, which is different from proposition 4.1 where the smallness of $K_0\|\omega\|$ is needed. This will allow us to pass through the approximation argument due to the perturbation of the initial guess.

We now derive a perturbation result on $\|x_n - \hat{x}_n\|$ and $\|T(x_n - \hat{x}_n)\|$ relative to the change of the initial guess. For simplicity of the presentation we set

$$\hat{e}_n := \hat{x}_n - x^\dagger, \qquad T_n = F'(x_n), \qquad \hat{T}_n = F'(\hat{x}_n).$$

It follows from (5.7) and (5.8) that

$$
\begin{aligned}
x_{n+1} - \hat{x}_{n+1} &= x_n - \hat{x}_n - g_{\alpha_n}(T_n^* T_n) T_n^* (F(x_n) - y) + g_{\alpha_n}(\hat{T}_n^* \hat{T}_n) \hat{T}_n^* (F(\hat{x}_n) - y) \\
&= r_{\alpha_n}(T^* T)(x_n - \hat{x}_n) - g_{\alpha_n}(T^* T) T^* \left(F(x_n) - F(\hat{x}_n) - T(x_n - \hat{x}_n) \right) \\
&\quad - \left[g_{\alpha_n}(T_n^* T_n) T_n^* - g_{\alpha_n}(T^* T) T^* \right] (F(x_n) - F(\hat{x}_n)) \\
&\quad - \left[g_{\alpha_n}(T_n^* T_n) T_n^* - g_{\alpha_n}(\hat{T}_n^* \hat{T}_n) \hat{T}_n^* \right] (F(\hat{x}_n) - y) .
\end{aligned}
$$

By telescoping this identity we obtain

$$
\begin{aligned}
x_{n+1} - \hat{x}_{n+1} = {} & \prod_{k=0}^{n} r_{\alpha_k}(T^* T)(x_0 - \hat{x}_0) \\
& - \sum_{j=0}^{n} \prod_{k=j+1}^{n} r_{\alpha_k}(T^* T) g_{\alpha_j}(T^* T) T^* \left(F(x_j) - F(\hat{x}_j) - T(x_j - \hat{x}_j) \right) \\
& - \sum_{j=0}^{n} \prod_{k=j+1}^{n} r_{\alpha_k}(T^* T) \left[g_{\alpha_j}(T_j^* T_j) T_j^* - g_{\alpha_j}(T^* T) T^* \right] (F(x_j) - F(\hat{x}_j)) \\
& - \sum_{j=0}^{n} \prod_{k=j+1}^{n} r_{\alpha_k}(T^* T) \left[g_{\alpha_j}(T_j^* T_j) T_j^* - g_{\alpha_j}(\hat{T}_j^* \hat{T}_j) \hat{T}_j^* \right] (F(\hat{x}_j) - y) .
\end{aligned}
$$
(5.13)

Lemma 5.4. *Let F satisfy Assumptions 3, let $\{g_\alpha\}$ satisfy Assumptions 1 and 2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.12). If $K_0\|e_0\|$ and $K_0\|\hat{e}_0\|$ are suitably small, then*

$$
(5.14) \qquad \|x_n - \hat{x}_n\| \le 2\|x_0 - \hat{x}_0\| \quad and \quad \|T(x_n - \hat{x}_n)\| \le 2c_0 s_n^{-1/2} \|x_0 - \hat{x}_0\|
$$

for $n = 0, 1, \cdots$.

Proof. We will show (5.14) by induction. Since $\|T\| \le \sqrt{\alpha_0}$, (5.14) holds for $n = 0$. In the following we will assume that (5.14) holds for $0 \le n \le l$, and show that it is also true for $n = l + 1$.

In view of Assumption 2, Lemma 3.1 with $a = b = 0$, and Lemma 3.2 with $\mu = 0$, it follows from (5.13) that

$$
\begin{aligned}
\|x_{l+1} - \hat{x}_{l+1}\| \le {} & \|x_0 - \hat{x}_0\| \\
& + C \sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1/2} \|F(x_j) - F(\hat{x}_j) - T(x_j - \hat{x}_j)\| \\
& + C \sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1/2} K_0 \|e_j\| \|F(x_j) - F(\hat{x}_j)\| \\
& + C \sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1/2} K_0 \|x_j - \hat{x}_j\| \|F(\hat{x}_j) - y\| .
\end{aligned}
$$
(5.15)

Next we multiply (5.13) by T. By using Assumption 2, Lemma 3.1 with $a = 1/2$ and $b = 0$, and Lemma 3.2 with $\mu = 1/2$, we obtain

$$\|T(x_{l+1} - \hat{x}_{l+1})\| \le s_l^{-1/2}\|x_0 - \hat{x}_0\|$$

$$+ C\sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1}\|F(x_j) - F(\hat{x}_j) - T(x_j - \hat{x}_j)\|$$

$$+ C\sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1}K_0\|e_j\|\|F(x_j) - F(\hat{x}_j)\|$$

$$(5.16) \qquad + C\sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1}K_0\|x_j - \hat{x}_j\|\|F(\hat{x}_j) - y\|.$$

From Lemma 5.2 and Assumption 3 it follows that

$$(5.17) \qquad \|F(\hat{x}_j) - y\| \lesssim \|T\hat{e}_j\| + K_0\|\hat{e}_j\|\|T\hat{e}_j\| \lesssim s_j^{-1/2}\|\hat{e}_0\|.$$

Moreover, By using Assumption 3 we have

$$F(x_j) - F(\hat{x}_j) - T(x_j - \hat{x}_j) = \int_0^1 [F'(\hat{x}_j + t(x_j - \hat{x}_j)) - T](x_j - \hat{x}_j)dt$$

$$= \int_0^1 [R(\hat{x}_j + t(x_j - \hat{x}_j), x^\dagger) - I]T(x_j - \hat{x}_j)dt.$$

Consequently

$$\|F(x_j) - F(\hat{x}_j) - T(x_j - \hat{x}_j)\| \le \int_0^1 \|R(\hat{x}_j + t(x_j - \hat{x}_j), x^\dagger) - I\|\|T(x_j - \hat{x}_j)\|dt$$

$$\le \frac{1}{2}K_0(\|e_j\| + \|\hat{e}_j\|)\|T(x_j - \hat{x}_j)\|.$$

With the help of Lemma 5.2 it yields

$$(5.18) \qquad \|F(x_j) - F(\hat{x}_j) - T(x_j - \hat{x}_j)\| \le K_0(\|e_0\| + \|\hat{e}_0\|)\|T(x_j - \hat{x}_j)\|.$$

This in particular implies

$$(5.19) \qquad \|F(x_j) - F(\hat{x}_j)\| \le 2\|T(x_j - \hat{x}_j)\|.$$

By virtue of (5.17), (5.18), (5.19) and the induction hypotheses, we have from (5.15) and (5.16) that

$$\|x_{l+1} - \hat{x}_{l+1}\| \le \|x_0 - \hat{x}_0\|$$

$$(5.20) \qquad + CK_0(\|e_0\| + \|\hat{e}_0\|)\|x_0 - \hat{x}_0\|\sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1/2}s_j^{-1/2}$$

and

$$\|T(x_{l+1} - \hat{x}_{l+1})\| \le s_l^{-1/2}\|x_0 - \hat{x}_0\|$$

$$(5.21) \qquad + CK_0(\|e_0\| + \|\hat{e}_0\|)\|x_0 - \hat{x}_0\|\sum_{j=0}^{l} \frac{1}{\alpha_j}(s_l - s_{j-1})^{-1}s_j^{-1/2}.$$

With the help of Lemma 4.1 and (5.1) we can derive

$$\|x_{l+1} - \hat{x}_{l+1}\| \le (1 + CK_0(\|e_0\| + \|\hat{e}_0\|)) \|x_0 - \hat{x}_0\|$$
$$\le 2\|x_0 - \hat{x}_0\|$$

and

$$\|T(x_{l+1} - \hat{x}_{l+1})\| \le (1 + CK_0(\|e_0\| + \|\hat{e}_0\|)) s_l^{-1/2} \|x_0 - \hat{x}_0\|$$
$$\le 2c_0 s_{l+1}^{-1/2} \|x_0 - \hat{x}_0\|$$

if $K_0\|e_0\|$ and $K_0\|\hat{e}_0\|$ are suitably small. The proof is thus complete. $\square$

Theorem 5.5. *Let F satisfy Assumptions 3, let $\{g_\alpha\}$ satisfy Assumptions 1 and 2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.12). If $e_0 \in \mathcal{N}(T)^\perp$ and $K_0\|e_0\|$ is suitably small, then*

(5.22) $$\lim_{n\to\infty} \|x_n - x^\dagger\| = 0 \quad and \quad \lim_{n\to\infty} s_n^{1/2}\|T(x_n - x^\dagger)\| = 0$$

for the sequence $\{x_n\}$ defined by (5.7).

Proof. Let $0 < \varepsilon < \|e_0\|$ be an arbitrarily small number. Since $e_0 \in \mathcal{N}(T)^\perp = \overline{\mathcal{R}(T^*)}$, there is an $\hat{x}_0 \in X$ such that $\hat{e}_0 := \hat{x}_0 - x^\dagger \in \mathcal{R}(T^*)$ and $\|x_0 - \hat{x}_0\| < \varepsilon$. Note that $K_0\|\hat{e}_0\| \le 2K_0\|e_0\|$. Thus, if $K_0\|e_0\|$ is suitably small, then for the sequence $\{\hat{x}_n\}$ defined by (5.8), it follows from Lemma 5.4 that

$$\|x_n - \hat{x}_n\| \le 2\|x_0 - \hat{x}_0\| < 2\varepsilon$$

and

$$s_n^{1/2}\|T(x_n - \hat{x}_n)\| \le 2c_0\|x_0 - \hat{x}_0\| < 2c_0\varepsilon$$

for all $n \ge 0$. On the other hand, since $\hat{e}_0 \in \mathcal{R}(T^*) = \mathcal{R}((T^*T)^{1/2}) \subset \mathcal{R}((T^*T)^{1/4})$, from Lemma 5.3 we have $\|\hat{e}_n\| \to 0$ and $s_n^{1/2}\|T\hat{e}_n\| \to 0$ as $n \to \infty$. Thus, there is a n_0 such that $\|\hat{e}_n\| < \varepsilon$ and $s_n^{1/2}\|T\hat{e}_n\| < c_0\varepsilon$ for all $n \ge n_0$. Consequently

$$\|e_n\| \le \|x_n - \hat{x}_n\| + \|\hat{e}_n\| < 3\varepsilon$$

and

$$s_n^{1/2}\|Te_n\| \le s_n^{1/2}\|T(x_n - \hat{x}_n)\| + s_n^{1/2}\|T\hat{e}_n\| < 3c_0\varepsilon$$

for all $n \ge n_0$. Since $\varepsilon > 0$ is arbitrarily small, we therefore obtain (5.22). $\square$

5.2. **Stability estimates.** In this subsection we will derive the stability estimates on $\|x_n^\delta - x_n\|$ for $0 \le n \le \hat{n}_\delta$, where $\hat{n}_\delta$ is defined by (5.2). We will use the notations

$$\mathcal{A} := F'(x^\dagger)^* F'(x^\dagger), \quad \mathcal{A}_n := F'(x_n)^* F'(x_n), \quad \mathcal{A}_n^\delta := F'(x_n^\delta)^* F'(x_n^\delta).$$

The main result is as follows.

Proposition 5.1. *Let F satisfy Assumptions 3, let $\{g_\alpha\}$ satisfy Assumptions 1 and 2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.12). If $K_0\|e_0\|$ is suitably small, then*

(5.23) $$\|x_n^\delta - x_n\| \lesssim s_n^{1/2}\delta$$

and

(5.24) $$\|F(x_n^\delta) - F(x_n) - y^\delta + y\| \le (1 + CK_0\|e_0\|)\,\delta$$

for $0 \le n \le \hat{n}_\delta$.

Proof. We first show (5.23) by establishing

$$(5.25) \qquad \|x_n^\delta - x_n\| \le 2b_2 c_2 s_n^{1/2}\delta \qquad \text{and} \qquad \|T(x_n^\delta - x_n)\| \le 3\delta$$

for $0 \le n \le \hat{n}_\delta$, where b_2 and c_2 are the constants appearing in (2.6) and (4.13) respectively. It is clear that (5.25) is true for $n = 0$. Now we assume that (5.25) is true for all $0 \le n \le l$ for some $l < \hat{n}_\delta$ and show that it is also true for $n = l + 1$. We set

$$v_n := F(x_n^\delta) - F(x_n) - T(x_n^\delta - x_n),$$
$$w_n := F(x_n^\delta) - F(x_n) - y^\delta + y.$$

It then follows from the definition of $\{x_n^\delta\}$ and $\{x_n\}$ that

$$
\begin{aligned}
x_{n+1}^\delta - x_{n+1} &= x_n^\delta - x_n - g_{\alpha_n}(\mathcal{A}_n^\delta)F'(x_n^\delta)^*(F(x_n^\delta) - y^\delta) \\
&\quad + g_{\alpha_n}(\mathcal{A}_n)F'(x_n)^*(F(x_n) - y) \\
&= r_{\alpha_n}(\mathcal{A})(x_n^\delta - x_n) - g_{\alpha_n}(\mathcal{A})F'(x^\dagger)^*\left(v_n - y^\delta + y\right) \\
&\quad - \left[g_{\alpha_n}(\mathcal{A}_n)F'(x_n)^* - g_{\alpha_n}(\mathcal{A})F'(x^\dagger)^*\right]w_n \\
&\quad - \left[g_{\alpha_n}(\mathcal{A}_n^\delta)F'(x_n^\delta)^* - g_{\alpha_n}(\mathcal{A}_n)F'(x_n)^*\right]\left(F(x_n^\delta) - y^\delta\right).
\end{aligned}
$$

By telescoping the above equation and noting that $x_0^\delta = x_0$ we obtain

$$
\begin{aligned}
x_{l+1}^\delta - x_{l+1} &= \sum_{j=0}^{l}\prod_{k=j+1}^{l} r_{\alpha_k}(\mathcal{A})g_{\alpha_j}(\mathcal{A})F'(x^\dagger)^*\left(y^\delta - y - v_j\right) \\
&\quad - \sum_{j=0}^{l}\prod_{k=j+1}^{l} r_{\alpha_k}(\mathcal{A})\left[g_{\alpha_j}(\mathcal{A}_j^\delta)F'(x_j^\delta)^* - g_{\alpha_j}(\mathcal{A}_j)F'(x_j)^*\right]\left(F(x_j^\delta) - y^\delta\right) \\
(5.26)\quad &\quad - \sum_{j=0}^{l}\prod_{k=j+1}^{l} r_{\alpha_k}(\mathcal{A})\left[g_{\alpha_j}(\mathcal{A}_j)F'(x_j)^* - g_{\alpha_j}(\mathcal{A})F'(x^\dagger)^*\right]w_j.
\end{aligned}
$$

In view of Assumption 1, Lemma 3.1 and Lemma 3.2 it follows that

$$
\begin{aligned}
\|x_{l+1}^\delta - x_{l+1}\| &\le b_2 \sum_{j=0}^{l}\frac{1}{\alpha_j}(s_l - s_{j-1})^{-1/2}\left(\delta + \|v_j\|\right) \\
&\quad + C\sum_{j=0}^{l}\frac{1}{\alpha_j}(s_l - s_{j-1})^{-1/2}K_0\|x_j^\delta - x_j\|\,\|F(x_j^\delta) - y^\delta\| \\
(5.27)\quad &\quad + C\sum_{j=0}^{l}\frac{1}{\alpha_j}(s_l - s_{j-1})^{-1/2}K_0\|e_j\|\|w_j\|.
\end{aligned}
$$

By multiplying (5.26) by T and using (4.3) we obtain with $\mathcal{B} = F'(x^\dagger)F'(x^\dagger)$ that

$$T(x_{l+1}^\delta - x_{l+1}) - y^\delta + y$$

$$= \prod_{j=0}^{l} r_{\alpha_j}(\mathcal{B})(y - y^\delta) - \sum_{j=0}^{l} \prod_{k=j+1}^{l} r_{\alpha_k}(\mathcal{B}) g_{\alpha_j}(\mathcal{B}) \mathcal{B} v_j$$

$$- \sum_{j=0}^{l} T \prod_{k=j+1}^{l} r_{\alpha_k}(\mathcal{A}) \left[g_{\alpha_j}(\mathcal{A}_j^\delta) F'(x_j^\delta)^* - g_{\alpha_j}(\mathcal{A}_j) F'(x_j)^* \right] (F(x_j^\delta) - y^\delta)$$

$$- \sum_{j=0}^{l} T \prod_{k=j+1}^{l} r_{\alpha_k}(\mathcal{A}) \left[g_{\alpha_j}(\mathcal{A}_j) F'(x_j)^* - g_{\alpha_j}(\mathcal{A}) F'(x^\dagger)^* \right] w_j.$$

It follows from Assumption 1, Lemma 3.1 and Lemma 3.2 that

$$\|T(x_{l+1}^\delta - x_{l+1}) - y^\delta + y\|$$

$$\leq \delta + C \sum_{j=0}^{l} \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} \|v_j\| + C \sum_{j=0}^{l} \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} K_0 \|e_j\| \|w_j\|$$

(5.28) $$+ C \sum_{j=0}^{l} \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} K_0 \|x_j^\delta - x_j\| \|F(x_j^\delta) - y^\delta\|.$$

With the help of Assumption 3, Lemma 5.1, Lemma 5.2 and the induction hypotheses we have

$$\|v_j\| \lesssim K_0 \left(\|e_j\| + \|e_j^\delta\| \right) \|T(x_j^\delta - x_j)\| \lesssim K_0 \|e_0\| \delta$$

and

$$\|w_j\| \leq \delta + \|T(x_j^\delta - x_j)\| + \|v_j\| \lesssim \delta.$$

Moreover, by using Assumption 3, Lemma 5.1 and (5.2) we have for $0 \leq j \leq l$

$$\|F(x_j^\delta) - y^\delta\| \leq \delta + \|T e_j^\delta\| + K_0 \|e_j^\delta\| \|T e_j^\delta\| \lesssim \delta + s_j^{-1/2} \|e_0\| \lesssim s_j^{-1/2} \|e_0\|.$$

Combining the above three inequalities with (5.27) and (5.28) and using the induction hypothesis $\|x_j^\delta - x_j\| \lesssim s_j^{1/2} \delta$ for $0 \leq j \leq l$ it follows that

(5.29) $$\|x_{l+1}^\delta - x_{l+1}\| \leq (b_2 + CK_0\|e_0\|) \delta \sum_{j=0}^{l} \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2}$$

and

(5.30) $$\|T(x_{l+1}^\delta - x_{l+1}) - y^\delta + y\| \leq \delta + K_0 \|e_0\| \delta \sum_{j=0}^{l} \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1}.$$

By Lemma 4.1 and the fact $s_l \leq s_{l+1}$ we obtain for small $K_0\|e_0\|$ that

$$\|x_{l+1}^\delta - x_{l+1}\| \leq (b_2 c_2 + CK_0\|e_0\|) s_l^{1/2} \delta \leq 2 b_2 c_2 s_{l+1}^{1/2} \delta$$

Moreover, with the help of (5.1) we can derive

$$\|T(x_{l+1}^\delta - x_{l+1}) - y^\delta + y\| \leq (1 + CK_0\|e_0\|) \delta.$$

Thus $\|T(x_{l+1}^\delta - x_{l+1})\| \leq 3\delta$ if $K_0\|e_0\|$ is suitably small. We therefore complete the proof of (5.25).

Next we will prove (5.24). From the above proof we in fact obtain

$$\|T(x_n^\delta - x_n) - y^\delta + y\| \le (1 + CK_0\|e_0\|)\delta, \qquad 0 \le n \le \hat{n}_\delta.$$

Therefore, it follows from Assumption 3, Lemma 5.1 and Lemma 5.2 that

$$\begin{aligned}
\|F(x_n^\delta) &- F(x_n) - y^\delta + y\| \\
&\le \|F(x_n^\delta) - F(x_n) - T(x_n^\delta - x_n)\| + \|T(x_n^\delta - x_n) - y^\delta + y\| \\
&\le K_0(\|e_n^\delta\| + \|e_n\|)\|T(x_n^\delta - x_n)\| + (1 + CK_0\|e_0\|)\delta \\
&\le (1 + CK_0\|e_0\|)\delta.
\end{aligned}$$

The proof is thus complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

5.3. Completion of the proof of Theorem 2.2.
We have shown in Lemma 5.1 that $n_\delta \le \hat{n}_\delta$. Thus we may use the definition of n_δ and Proposition 5.1 to obtain

$$(5.31) \qquad \|F(x_{n_\delta}) - y\| \le \|F(x_{n_\delta}^\delta) - y^\delta\| + \|F(x_{n_\delta}^\delta) - F(x_{n_\delta}) - y^\delta + y\| \lesssim \delta$$

and for $0 \le n < n_\delta$

$$\begin{aligned}
\tau\delta \le \|F(x_n^\delta) - y^\delta\| &\le \|F(x_n^\delta) - F(x_n) - y^\delta + y\| + \|F(x_n) - y\| \\
&\le (1 + CK_0\|e_0\|)\delta + \|F(x_n) - y\|.
\end{aligned}$$

Since $\tau > 1$, if $K_0\|e_0\|$ is suitably small then

$$(5.32) \qquad\qquad\qquad \delta \lesssim \|F(x_n) - y\| \lesssim \|Te_n\|, \qquad 0 \le n < n_\delta.$$

We now prove the convergence of $x_{n_\delta}^\delta$ to $x^\dagger$ as $\delta \to 0$. Assume first that there is a sequence $\delta_k \searrow 0$ such that $n_k := n_{\delta_k} \to n$ as $k \to \infty$ for some finite integer n. Without loss of generality, we can assume that $n_k = n$ for all k. It then follows from (5.31) that $F(x_n) = y$. Thus, from (5.7) we can conclude that $x_j = x_n$ for all $j \ge n$. Since Theorem 5.5 implies $x_j \to x^\dagger$ as $j \to \infty$, we must have $x_n = x^\dagger$, which together with Proposition 5.1 implies $x_{n_k}^{\delta_k} \to x^\dagger$ as $k \to \infty$.

Assume next that there is a sequence $\delta_k \searrow 0$ such that $n_k := n_{\delta_k} \to \infty$ as $k \to \infty$. Then Theorem 5.5 and (5.32) imply that $\|e_{n_k}\| \to 0$ and $s_{n_k}^{1/2}\delta_k \to 0$ as $k \to \infty$. Consequently, by Proposition 5.1 we again obtain $x_{n_k}^{\delta_k} \to x^\dagger$ as $k \to \infty$.

6. Numerical results

In this section we present some numerical results to test the theoretical conclusions given in Theorems 2.1 and 2.2. We consider the estimation of the coefficient c in the two-point boundary value problem

$$(6.1) \qquad\qquad\qquad \begin{cases} -u'' + cu = f & \text{in } (0,1) \\ u(0) = g_0, & u(1) = g_1 \end{cases}$$

from the L^2 measurement u^δ of the state variable u, where g_0, g_1 and $f \in L^2[0,1]$ are given. This inverse problem reduces to solving (1.1) with the nonlinear operator $F : D(F) \subset L^2[0,1] \mapsto L^2[0,1]$ defined as the parameter-to-solution mapping $F(c) := u(c)$, where $u(c)$ denotes the unique solution of (6.1). It is well known that F is well-defined on

$$D(F) := \left\{ c \in L^2[0,1] : \|c - \hat{c}\|_{L^2} \le \gamma \text{ for some } \hat{c} \ge 0 \text{ a.e.} \right\}$$

with some $\gamma > 0$. Moreover, F is Fréchet differentiable, the Fréchet derivative and its adjoint are given by

$$F'(c)h = -A(c)^{-1}(hu(c)),$$
$$F'(c)^*w = -u(c)A(c)^{-1}w,$$

where $A(c) : H^2 \cap H_0^1 \mapsto L^2$ is defined by $A(c)u = -u'' + cu$. It has been shown in [5] that if, for the sought solution $c^\dagger$, $|u(c^\dagger)| \geq \kappa > 0$ on $[0,1]$, then Assumption 3 is satisfied in a neighborhood of $c^\dagger$.

In the following we report some numerical results on the method given by (1.5) and (1.6) with g_α defined by (2.14), which, in the current context, defines the iterative solutions $\{c_n^\delta\}$ by

$$u_{n,0} = c_n^\delta,$$
$$u_{n,l+1} = u_{n,l} - F'(c_n^\delta)^* \left(F(c_n^\delta) - u^\delta - F'(c_n^\delta)(c_n^\delta - u_{n,l}) \right), \quad 0 \leq l \leq [1/\alpha_n] - 1,$$
$$c_{n+1}^\delta = u_{n,[1/\alpha_n]}.$$

and determines the stopping index n_δ by

$$(6.2) \qquad \|F(c_{n_\delta}^\delta) - u^\delta\| \leq \tau\delta < \|F(c_n^\delta) - u^\delta\|, \quad 0 \leq n < n_\delta.$$

During the computation, all differential equations are solved approximately by finite difference method by dividing the interval $[0,1]$ into $m+1$ subintervals with equal length $h = 1/(m+1)$; we take $m = 100$ in our actual computation.

Example 1. We estimate c in (6.1) by assuming $f(t) = (1+t)(1+t-0.8\sin(2\pi t))$, $g_0 = 1$ and $g_1 = 2$. If $u(c^\dagger) = 1 + t$, then $c^\dagger = 1 + t - 0.8\sin(2\pi t)$ is the sought solution. When applying the above method, we take $\alpha_n = 2^{-n}$ and use random noise data u^δ satisfying $\|u^\delta - u(c^\dagger)\|_{L^2[0,1]} = \delta$ with noise level $\delta > 0$. As an initial guess we choose $c_0 = 1 + t$. One can show that $c_0 - c^\dagger \in \mathcal{R}(F'(c^\dagger)^*)$. Thus, according to Theorem 2.1, the expected rate of convergence should be $O(\delta^{1/2})$.

Table 1. Numerical results for Example 1 with $\alpha_n = 2^{-n}$ and three distinct values of τ, where n_δ is determined by (6.2), $error := \|c_{n_\delta}^\delta - c^\dagger\|_{L^2}$, and

$$ratio := error/\delta^{1/2}$$

	$\tau = 1.1$			$\tau = 2.0$			$\tau = 4.0$		
δ	k_δ	$error$	$ratio$	k_δ	$error$	$ratio$	k_δ	$error$	$ratio$
10^{-2}	12	$4.67e-2$	0.47	9	$2.90e-1$	2.90	1	$5.65e-1$	5.65
10^{-3}	14	$1.47e-2$	0.47	12	$3.89e-2$	1.23	12	$3.89e-2$	1.23
10^{-4}	16	$4.30e-3$	0.43	15	$5.30e-3$	0.53	14	$8.70e-3$	0.87
10^{-5}	18	$1.30e-3$	0.40	17	$1.80e-3$	0.56	16	$2.80e-3$	0.87
10^{-6}	21	$4.45e-4$	0.44	19	$6.41e-4$	0.64	18	$1.00e-3$	1.03

The numerical result is reported in Table 1. In order to see the effect of τ in the discrepancy principle (6.2), we consider the three distinct values $\tau = 1.1$, 2 and 4. In order to indicate the dependence of the convergence rates on the noise level, different values of δ are selected. The rates in Table 1 coincide with Theorem 2.1 very well. Table 1 indicates also that the absolute error increases with respect to τ. Thus, in numerical computation, one should use smaller τ if possible.

In order to visualize the computed solutions, we plot in Figure 1 the results obtained for $\tau = 1.1$ and various values of the noise level δ, where the solid, dashed,

and dash-dotted curves denote the exact solution $c^\dagger$, the initial guess c_0, and the computed solution $c_{n_\delta}^\delta$ respectively. It clearly indicates the fast convergence as $\delta \to 0$ as reported in Table 1.
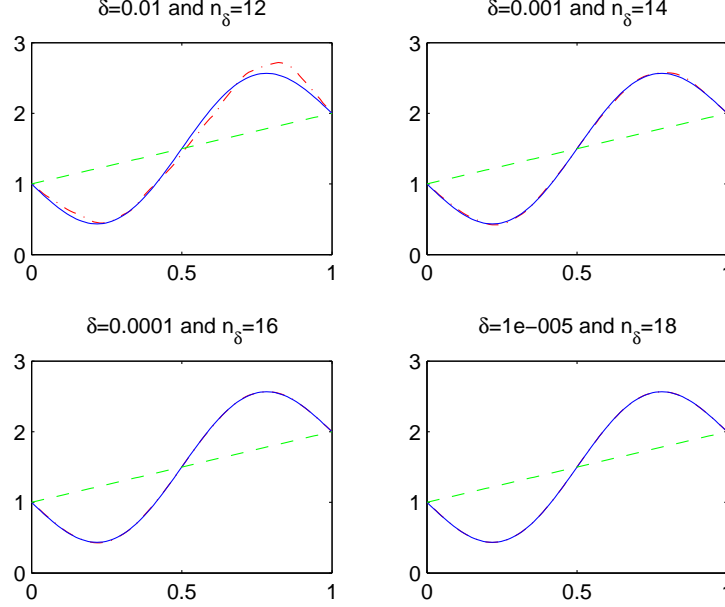

FIGURE 1. Comparison on the computed and exact solution for Example 1 with $\tau = 1.1$

Example 2. We repeat Example 1 but with $\tau = 1.1$ and the initial guess $c_0 = 2 - t$. Now $c_0 - c^\dagger \notin \mathcal{R}(F'(c^\dagger)^*)$, and in fact $c_0 - c^\dagger$ has no source-wise representation $c_0 - c^\dagger \in \mathcal{R}((F'(c^\dagger)^* F'(c^\dagger))^\nu)$ with a good $\nu > 0$. However, Theorem 2.2 asserts that $\|c_{n_\delta}^\delta - c^\dagger\|_{L^2[0,1]} \to 0$ as $\delta \to 0$. Figure 2 clearly indicates such convergence although the convergence speed could be quite slow which is typical for inverse problems.

REFERENCES

[1] A. B. Bakushinsky and M. Yu. Kokurin, *Iterative Methods for Approximate Solutions of Inverse Problems*, Mathematics and its applications, Springer, 2004.

[2] R. S. Dembo, S. C. Eisenstat and T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal., 19 (1982), 400–408.

[3] M. Hanke, *A regularizing Levenberg-Marquardt scheme with applications to inverse ground-water filtration problems*, Inverse Problems, 13(1997), 79–95.

[4] M. Hanke, *The regularizing Levenberg-Marquardt scheme is of optimal order*, J. Integeral Equations and Applications, 22 (2010), no. 2, 259–283.

[5] M. Hanke, A. Neubauer and O. Scherzer, *A convergence analysis of the Landweber iteration for nonlinear ill-posed problems*, Numer. Math., 72 (1995), 21–37.

[6] M. Hochbruck, M. Hönig and A. Ostermann, *A convergence analysis of the exponential Euler iteration for nonlinear ill-posed problems*, Inverse Problems, 25 (2009), no.7, article no. 075009.
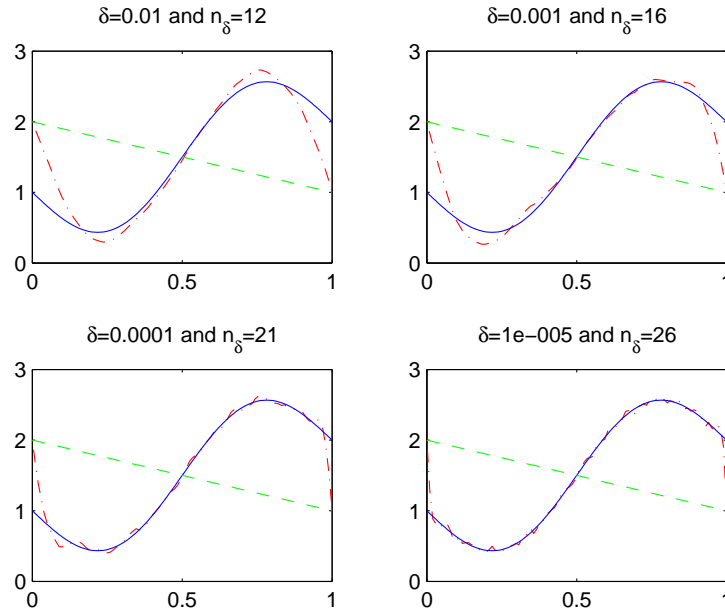
FIGURE 2. Comparison on the computed and exact solutions for Example 2 with $\tau = 1.1$

[7] Q. Jin, *On a regularized Levenberg-Marquardt method for solving nonlinear inverse problems*, Numer. Math., 115 (2010), no. 2, 229–259.

[8] Q. Jin and U. Tautenhahn, *Inexact Newton regularization methods in Hilbert scales*, Numer. Math., accepted for publication, 2010.

[9] A. Lechleiter and A. Rieder, *Towards a general convergence theory for inexact Newton regularizations*, Numer. Math. 114 (2010), no. 3, 521–548.

[10] P. Pornsawad and C. Böckmann, *Convergence rate analysis of the first-stage Runge-Kutta type regularizations*, Inverse Problems 26 (2010), no 3, article no. 035005.

[11] A. Rieder, *On the regularization of nonlinear ill-posed problems via inexact Newton iterations*, Inverse Problems, 15(1999), 309–327.

[12] A. Rieder, *On convergence rates of inexact Newton regularizations*, Numer. Math. 88(2001), 347–365.

DEPARTMENT OF MATHEMATICS, VIRGINIA TECH, BLACKSBURG, VA 24061

E-mail address: `qnjin@math.vt.edu`